

# Projeto Integrador IV - UNIVESP 2024

## Objetivo

Este relatório apresenta a análise de dados e técnicas de machine learning aplicadas aos dados fornecidos.

## Análise de Componentes Principais (PCA)

A PCA (Análise de Componentes Principais) é uma técnica de redução de dimensionalidade que transforma os dados em um novo sistema de coordenadas. As novas coordenadas, ou componentes principais, são combinações lineares das variáveis originais.

- Componente Principal 1:** Indica a direção da maior variância dos dados.
- Componente Principal 2:** Indica a segunda maior direção de variância.

## Resultados da Regressão Linear

A regressão linear foi utilizada para prever a variável dependente com base nas variáveis independentes. O gráfico mostra a relação entre os valores reais e previstos pela regressão linear.

## Distribuição das Variáveis

- Histograma:** O histograma da primeira coluna numérica mostra a distribuição dos dados, ajudando a entender a frequência das diferentes faixas de valores.
- Boxplot:** O boxplot exibe a distribuição das variáveis numéricas, ajudando a identificar outliers e a dispersão dos dados.
- Gráfico de Barras:** O gráfico de barras mostra a média das variáveis numéricas, destacando quais variáveis apresentam maiores médias e tendências gerais.